

#2
Rieko Furukawa
4/1/01

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

IN RE APPLICATION OF: Rieko FURUKAWA, et al.

GAU:

SERIAL NO: New Application

EXAMINER:

FILED: Herewith

FOR: VIDEO ENCODING APPARATUS AND METHOD

1c966 U.S. PTO
09/769309



REQUEST FOR PRIORITY

ASSISTANT COMMISSIONER FOR PATENTS
WASHINGTON, D.C. 20231

SIR:

- ☐ Full benefit of the filing date of U.S. Application Serial Number, filed, is claimed pursuant to the provisions of 35 U.S.C. §120.
- ☐ Full benefit of the filing date of U.S. Provisional Application Serial Number, filed, is claimed pursuant to the provisions of 35 U.S.C. §119(e).
- ☒ Applicants claim any right to priority from any earlier filed applications to which they may be entitled pursuant to the provisions of 35 U.S.C. §119, as noted below.

In the matter of the above-identified application for patent, notice is hereby given that the applicants claim as priority:

<u>COUNTRY</u>	<u>APPLICATION NUMBER</u>	<u>MONTH/DAY/YEAR</u>
Japan	2000-053823	February 29, 2000

Certified copies of the corresponding Convention Application(s)

- ☒ are submitted herewith
- ☐ will be submitted prior to payment of the Final Fee
- ☐ were filed in prior application Serial No. filed
- ☐ were submitted to the International Bureau in PCT Application Number .
Receipt of the certified copies by the International Bureau in a timely manner under PCT Rule 17.1(a) has been acknowledged as evidenced by the attached PCT/IB/304.
- ☐ (A) Application Serial No.(s) were filed in prior application Serial No. filed ; and
(B) Application Serial No.(s)
 - ☐ are submitted herewith
 - ☐ will be submitted prior to payment of the Final Fee

Respectfully Submitted,

OBLON, SPIVAK, McCLELLAND,
MAIER & NEUSTADT, P.C.

Marvin J. Spivak

Registration No. 24 913

C. Irvin McClelland

Registration Number 21,124



22850

日本国特許庁
PATENT OFFICE
JAPANESE GOVERNMENT

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日
Date of Application:

2000年 2月29日

出願番号
Application Number:

特願2000-053823

出願人
Applicant(s):

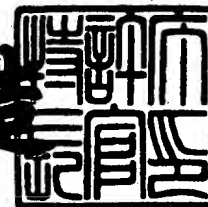
株式会社東芝

CERTIFIED COPY OF
PRIORITY DOCUMENT

2000年10月27日

特許庁長官
Commissioner,
Patent Office

及川耕造



出証番号 出証特2000-3087831

【書類名】 特許願

【整理番号】 A000000471

【提出日】 平成12年 2月29日

【あて先】 特許庁長官 殿

【国際特許分類】 H04N 7/30

【発明の名称】 動画像符号化装置および動画像符号化方法

【請求項の数】 10

【発明者】

【住所又は居所】 神奈川県川崎市幸区小向東芝町1番地 株式会社東芝研究開発センター内

【氏名】 古川 理恵子

【発明者】

【住所又は居所】 神奈川県川崎市幸区小向東芝町1番地 株式会社東芝研究開発センター内

【氏名】 菊池 義浩

【特許出願人】

【識別番号】 000003078

【氏名又は名称】 株式会社 東芝

【代理人】

【識別番号】 100058479

【弁理士】

【氏名又は名称】 鈴江 武彦

【電話番号】 03-3502-3181

【選任した代理人】

【識別番号】 100084618

【弁理士】

【氏名又は名称】 村松 貞男

【選任した代理人】

【識別番号】 100068814

【弁理士】

【氏名又は名称】 坪井 淳

【選任した代理人】

【識別番号】 100092196

【弁理士】

【氏名又は名称】 橋本 良郎

【選任した代理人】

【識別番号】 100091351

【弁理士】

【氏名又は名称】 河野 哲

【選任した代理人】

【識別番号】 100088683

【弁理士】

【氏名又は名称】 中村 誠

【選任した代理人】

【識別番号】 100070437

【弁理士】

【氏名又は名称】 河井 将次

【手数料の表示】

【予納台帳番号】 011567

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【ブルーフの要否】 要

【書類名】 明細書

【発明の名称】 動画像符号化装置および動画像符号化方法

【特許請求の範囲】

【請求項 1】 入力画像信号を時間的に連続した少なくとも一つのフレームからなるシーンに分割して各シーン毎に統計的特徴量を算出する特徴量算出手段と、

前記特徴量算出手段により算出された統計的特徴量に基づいて前記各シーン毎の符号化パラメタを生成する符号化パラメタ生成手段と、

前記符号化パラメタ生成手段で生成された符号化パラメタで前記入力画像信号を符号化した際の発生符号量を予測する発生符号量予測手段と、

前記発生符号量予測手段による発生符号量の予測結果に基づき前記符号化パラメタを修正する符号化パラメタ修正手段と、

前記修正された符号化パラメタを用いて前記入力動画像信号を符号化して符号化列を生成する符号化手段と、

前記符号化手段により生成された符号化列を符号化出力として出力する出力手段とを具備することを特徴とする動画像符号化装置。

【請求項 2】 前記特徴量算出手段は、動きベクトルに関する統計的特徴量に基づいて複数のシーンそれぞれをあらかじめ決められた複数のシーンタイプに分類し、

前記符号化パラメタ生成手段は、前記シーンタイプの分類結果に従って前記各シーン毎にフレームレートと量子化幅の重み付けを設定することを特徴とする請求項 1 記載の動画像符号化装置。

【請求項 3】 前記符号化パラメタ修正手段は、ユーザが指定した目標ビットレートを満たすように、前記入力画像信号を符号化するためのシーン毎のビットレートを符号化パラメタとして修正し算出することを特徴とする請求項 1 記載の動画像符号化装置。

【請求項 4】 前記符号化手段は、前記入力画像信号のシーン毎に指定されたビットレート情報・フレームレート情報を前記修正された符号化パラメタとして受け取り、シーン毎のビットレートが指定された値になるように前記ビットレ

ート情報・フレームレート情報に従って前記入力画像信号を符号化することを特徴とする請求項 1 記載の動画像符号化装置。

【請求項 5】 前記符号化手段は、前記入力画像信号のシーン毎に指定されたビットレート情報を前記修正された符号化パラメタとして受け取り、前記ビットレート情報に従ってシーン毎のビットレートが指定された値に変化する仮想バッファの占有量を用いて、量子化ステップサイズと符号化を行うフレームの間隔を決定することを特徴とする請求項 1 記載の動画像符号化装置。

【請求項 6】 前記符号化パラメタ生成手段は、前記特徴量算出手段により算出されたマクロブロック毎の輝度の分布に関する統計的特徴量に基づいて、前記各シーンを構成する符号化対象フレームのマクロブロックに対して量子化幅の重み付けを設定することを特徴とする請求項 1 または 2 記載の動画像符号化装置。

【請求項 7】 入力画像信号を時間的に連続した少なくとも一つのフレームからなるシーンに分割して各シーン毎に統計的特徴量を算出する特徴量算出ステップと、

前記特徴量算出ステップにより算出された統計的特徴量に基づいて前記各シーン毎の符号化パラメタを生成する符号化パラメタ生成ステップと、

前記符号化パラメタ生成ステップで生成された符号化パラメタで前記入力画像信号を符号化した際の発生符号量を予測する発生符号量予測ステップと、

前記発生符号量予測ステップによる発生符号量の予測結果に基づき前記符号化パラメタを修正する符号化パラメタ修正ステップと、

前記修正された符号化パラメタを用いて前記入力動画像信号を符号化して符号化列を生成する符号化ステップと、

前記符号化ステップにより生成された符号化列を符号化出力として出力するステップとを具備することを特徴とする動画像符号化方法。

【請求項 8】 前記特徴量算出ステップは、動きベクトルに関する統計的特徴量に基づいて複数のシーンそれぞれをあらかじめ決められた複数のシーンタイプに分類し、

前記符号化パラメタ生成手段は、前記シーンタイプの分類結果に従って前記各

シーン毎にフレームレートと量子化幅の重み付けを設定することを特徴とする請求項 7 記載の動画像符号化方法。

【請求項 9】 前記符号化パラメタ生成ステップは、前記特徴量算出ステップにより算出されたマクロブロック毎の輝度の分布に関する統計的特徴量に基づいて、前記各シーンを構成する符号化対象フレームのマクロブロックに対して量子化幅の重み付けを設定することを特徴とする請求項 7 または 8 記載の動画像符号化方法。

【請求項 10】 入力画像信号を符号化するためのコンピュータプログラムが記録された記録媒体であって、

前記コンピュータプログラムは、

入力画像信号を時間的に連続した少なくとも一つのフレームからなるシーンに分割して各シーン毎に統計的特徴量を算出する特徴量算出ステップと、

前記特徴量算出ステップにより算出された統計的特徴量に基づいて前記各シーン毎の符号化パラメタを生成する符号化パラメタ生成ステップと、

前記符号化パラメタ生成ステップで生成された符号化パラメタで前記入力画像信号を符号化した際の発生符号量を予測する発生符号量予測ステップと、

前記発生符号量予測ステップによる発生符号量の予測結果に基づき前記符号化パラメタを修正する符号化パラメタ修正ステップと、

前記修正された符号化パラメタを用いて前記入力動画像信号を符号化して符号化列を生成する符号化ステップと、

前記符号化ステップにより生成された符号化列を符号化出力として出力するステップとを具備することを特徴とする記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、インターネットなどによる動画像伝送システムや画像データベースシステムに使用される動画像符号化装置および動画像符号化方法に関し、特に 2 パス符号化という手法でシーンの内容に応じた符号化パラメタに従って符号化を行い、データサイズを大きくすることなくシーン毎にまとまりのある見やすい復

号画像を提供できるように改善された動画像符号化装置および動画像符号化方法に関する。

【 0 0 0 2 】

【従来の技術】

動画像の圧縮に用いられるMPEG方式は、動画像データのフレームどうしでの動き補償後の残差をDCT (Discrete Cosine Transform: 離散コサイン変換) 処理し、その係数を量子化することでデータを圧縮するものである。

【 0 0 0 3 】

MPEG方式に基づく従来の動画像符号化方式では、圧縮動画像データを伝送レートが規定されている伝送路によって伝送したり、記録容量に限りがある蓄積媒体に記録するために、出力される符号化ビットストリームのビットレートが指定された値となるようにフレームレートや量子化幅などの符号化パラメタを設定して符号化を行うというレート制御と呼ばれる処理が行われている。

【 0 0 0 4 】

多くのレート制御では、前フレームの発生符号量に応じて次のフレームまでの間隔と次フレームの量子化幅を決める方法がとられている。そのため、画面の動きが大きいシーンになると発生符号量が増えるので画質が急に劣化する。図10 (a) は従来のレート制御について示す図である。従来のレート制御では401のように一定の目標ビットレートと403のように一定のフレームレートが設定されている。また、実際のビットレートを402、実際のフレームレートを404で表す。

【 0 0 0 5 】

従来のレート制御では、あらかじめ設定されたフレームスキップしきい値のバッファサイズと現時点でのバッファレベルとの差 (余裕度) を基にフレームレートを決め、現時点でのバッファがしきい値よりも小さい時には一定のフレームレートで符号化を行い、現時点でのバッファがしきい値を上回るとフレームレートを下げるように制御する。したがって、動きの激しいシーンに切り替わると、発生符号量が急増するために図11のようなフレームスキップがおり404のよ

うにフレームレートが落ちてしまう。

【0006】

このように、従来のレート制御は画像内容とは無関係に次フレームの符号量を定めているため、画面の動きが大きくなるシーンでフレーム間隔が広くなり過ぎて動きが不自然になったり、量子化幅が適切でないために画像が歪んで見づらさを感じさせることがあった。

【0007】

一方、2パス符号化と呼ばれる方法によりレート制御を行う方式も知られている。しかし、多くの手法では符号量の変化のみに着目しており、画像内容と符号量の関係を考慮したものは、例えばシェードイン・シェードアウトなど（特開平10-336641号公報）の特別な場合に限られていた。

【0008】

【発明が解決しようとする課題】

上述のように、従来の動画像符号化装置では、画像の内容に無関係にフレームレートや量子化幅が決められていたためにオブジェクトの動きが激しいシーンにおいて急激にフレームレートが落ちたり、量子化幅が適切でないために画像が歪むなど、画質劣化が目立ちやすいという問題点があった。

【0009】

そこで、本発明は符号化する動画像の内容に適した効率的な符号配分を行えるようにし、視覚的にまとまりのある見やすい画像を生成することが可能な動画像符号化装置および動画像符号化方法を提供することを目的とする。

【0010】

【課題を解決するための手段】

上述の課題を解決するため、本発明はまず入力動画像信号を時間的に連続した少なくとも一つのフレームからなるシーンに分割して、各シーン毎に統計的特徴量を算出し、この統計的特徴量に基づいてシーンの内容を推定する。さらに、目標ビットレートに対してバッファにある程度の余裕をもたせることができる場合、シーン内容に応じて符号化ビットレートをユーザが設定した上限・下限内で配分し、平均ビットレートがあらかじめ指定されたビットレートを満たすような効

率の良い符号化パラメタを決定する。この符号化パラメタを用いて入力画像信号を符号化し、同じデータサイズでも視覚的に見やすい復号画像を得ることを基本的な特徴とする。

【0011】

ここで、統計的特徴量は、例えば入力動画像信号の各フレーム内に存在する動きベクトルや輝度値をシーン毎に集計することにより算出される。加えて、特徴量から入力動画像信号を得るときに使用したカメラの動きおよび画像内のオブジェクトの動きを推定した結果を用いて、シーンそれぞれをあらかじめ決められた複数のシーンタイプに分類し、その分類結果をフレームレートや量子化幅の割り当てに反映させる。また、統計的特徴量としてマクロブロック毎に輝度値の分布を調べることで、モスキートノイズが発生しやすいマクロブロックやオブジェクトのエッジが存在するマクロブロックの量子化幅を他のマクロブロックに比して相対的に小さくし、画質の向上を図る。

【0012】

本発明の2パス目における符号化では、前記算出されたシーン毎に適切なビットレートとフレームレートを与えることで、シーン毎のレート制御が行われ、従来のレート制御機構を大きく変更することなくシーン内容に応じた符号化が行える。

【0013】

上記のような2パスの手法を用いて、目標符号量と同じデータサイズで良好な復号画像を得る符号化を実現することができる。

【0014】

【発明の実施の形態】

以下、図面を参照して本発明の実施の形態を説明する。

【0015】

図1は本発明の一実施形態に係る動画像符号化装置の構成を示すブロック図である。入力動画像信号100は、同じ信号を繰り返し複数回再生可能なデジタルVTRやDVDシステムのようなビデオ記録再生装置によって再生される動画像信号（映像信号）であり、画像特徴量算出部31と符号化部のフレームメモリ

1 1 に入力される。

【 0 0 1 6 】

M P E G 符号化で用いられている動き補償適応予測・離散コサイン変換符号化・量子化方式については文献（『M P E G』テレビジョン学会 編、オーム社）に詳しく述べられているので、ここでは動作の概略のみを記し、従来方式との差異を詳細に説明する。

【 0 0 1 7 】

図 1 に示した符号化装置では、動画像信号 1 0 0 をフレームメモリ 1 1 に入力する前に画像特徴量算出部 3 1 に入力し、符号化パラメタを算出し（1 パス目）、その算出された符号化パラメタ 1 3 4 と動画像信号 1 0 0 とを入力として、符号化ビットストリーム 2 0 0 を出力する（2 パス目）。図 2 は符号化の流れを示すフローチャートである。

【 0 0 1 8 】

1 パス目の処理では、まず動画像信号 1 0 0 が画像特徴量算出部 3 1 に入力され、画像特徴量算出部 3 1 にて動画像信号 1 0 0 のシーン分割と各フレームについての画像特徴量の計算が行われる（ステップ S 1 1, S 1 2）。各シーンは、時間的に連続した少なくとも一つのフレームからなる。

【 0 0 1 9 】

画像特徴量は、例えば動きベクトルの数、分布、ノルムの大きさ、動き補償後の残差、輝度・色差の分散等である。算出された特徴量は前記分割されたシーン毎に集計され、これにより各シーン毎に統計的特徴量が算出される。画像特徴量算出部 3 1 によってシーン毎に得られた統計的特徴量 1 3 0 は符号化パラメタ生成部 3 2 に送られ、そこで各シーン毎に適切な符号化パラメタが生成される（ステップ S 1 3）。この場合、統計的特徴量 1 3 0 に基づいて、入力動画像信号を得るときに使用したカメラの動きおよび画像内のオブジェクトの動きなどが各シーン毎に推定され、その推定結果を基に、各シーンに適切なフレームレートと量子化幅の算出が行われる。また、マクロブロック毎に輝度値の分布を調べることで、モスキートノイズが発生しやすいマクロブロックやオブジェクトのエッジが存在するマクロブロックの量子化幅を他のマクロブロックに比して相対的に小さ

くするといった、マクロブロックの量子化幅の設定が行われる。

【0020】

符号化パラメタ生成部32で得られた符号化パラメタ131は発生符号量予測部33に入力される。この発生符号量予測部33では、符号化パラメタとして算出されたフレームレートと量子化幅で符号化を行ったときの発生符号量をシーン毎に算出することにより、動画像信号100を符号化した際の発生符号量が予測される（ステップS14）。予測値132は符号化パラメタ修正部34に送られる。

【0021】

予測された発生符号量がユーザが設定した目標符号量133と大きく異なる場合には、予測した符号量がユーザが設定した符号量を満たすように符号化パラメタ修正部34においてパラメタの修正を行う（ステップS15, S16）。このようにして、2パス目で用いられる符号化パラメタである、シーン毎のビットレートとフレームレート134が算出される（ステップS17）。

【0022】

2パス目では、算出されたシーン毎のフレームレートとビットレート134を用いて、入力動画像信号100に対する符号化が符号化部にて行われ（ステップS18）、シーンの内容に応じて適切に符号量が配分されたビットストリーム200が出力される（ステップS19）。

【0023】

ここで、符号化部における符号化処理を簡単に説明する。

【0024】

フレームメモリ11に入力された入力動画像信号100は、マクロブロック単位に分割されて減算器12に入力され、ここで後述するようにして生成される予測画像信号との差分がとられて、予測残差信号が生成される。この予測残差信号はDCT（離散コサイン変換）部13により離散コサイン変換される。DCT部13で得られたDCT係数データは、量子化部14で量子化される。量子化された信号は2分岐され、一方は可変長符号化部20で動きベクトル情報などと共に可変長符号化され、そしてバッファ21で送信レートが平滑化されて符号化デー

タ（ビットストリーム）として出力される。ここで、バッファ21はビットレート情報134に従ってシーン毎のビットレートが指定された値に変化する仮想バッファとして用いられ、量子化ステップサイズと符号化を行うフレームの間隔はその仮想バッファの占有量に基づいて制御される。

【0025】

一方、2分岐された信号の他方は、逆量子化部15及び逆DCT（逆離散コサイン変換）部16により、量子化部14およびDCT部13の処理と逆の処理を順次受けた後、加算器17で予測画像信号と加算されることにより、局部復号信号が生成される。この局部復号信号は、ビデオメモリ18に蓄えられ、動き補償予測部19に入力される。動き補償予測部19では、入力動画像信号とビデオメモリ18に蓄えられた前フレームの画像との相関から動き検出、動きベクトルの探索などの動き補償処理が行われて、予測画像信号が生成される。

【0026】

次に、本実施形態の特徴部分である画像特徴量算出部31の個々の処理についてさらに詳しく説明する。

【0027】

<画像特徴量算出部・シーン分割>

入力動画像信号100は、隣接するフレーム間の差分によりフラッシュフレームや雑音フレームなどのフレームを除いて複数のシーンに分割される。ここで、フラッシュフレームとは例えばニュース番組でのインタビューシーンで、フラッシュ（ストロボ）が発光した瞬間のように、輝度が急激に高くなるフレームである。また、雑音フレームとはカメラの振れ等により画像が大きく劣化したフレームである。

【0028】

例えば、シーン分割は以下の様に行う。図3（a）のように隣接した*i*番目のフレームと（*i*+1）番目のフレームの間の差分値があらかじめ定められたしきい値を超え、かつ*i*番目のフレームと（*i*+2）番目のフレームの間の差分値も同様にしきい値を超えているならば、（*i*+1）番目のフレームはシーンの区切りと判定する。一方、図3（b）のように、*i*番目のフレームと（*i*+1）番目

のフレームの間の差分値があらかじめ定められたしきい値を超えていても、 i 番目のフレームと $(i + 2)$ 番目のフレームの間の差分値がしきい値を超えていなければ、 $(i + 1)$ 番目のフレームはシーンの区切りとしない。

【0029】

＜画像特徴量算出部・動きベクトルの計算＞

シーン分割の他に、入力動画像信号 100 の全フレームについて、フレーム内のマクロブロックの動きベクトルと動き補償残差、輝度値の平均・分散等を計算する。なお、特徴量の計算は全フレームに対してでもよいし、画像の性質を解析できる範囲で数フレームおきに計算してもよい。

【0030】

i 番目のフレームについての動領域のマクロブロックの数、動き補償残差、輝度値の分散をそれぞれ $MvNum(i)$ 、 $MeSad(i)$ 、 $Yvar(i)$ とする。ここで、動領域とは 1 フレーム中で前フレームからの動きベクトル $\neq 0$ であるマクロブロックの領域を指す。 j 番目のシーンに対して、そのシーンに含まれるフレームすべての $MvNum(i)$ 、 $MeSad(i)$ 、 $Yvar(i)$ の平均値をそれぞれ $Mvnum_j$ 、 $MeSad_j$ 、 $Yvar_j$ とし、それらを j 番目のシーンの特徴量の代表値とする。

【0031】

＜画像特徴量算出部・シーンの分類＞

さらに本実施形態では、動きベクトルを用いて次のようなシーンの分類を行いシーン内容を推定する。

【0032】

各々のフレームに対する動きベクトルを算出した後、動きベクトルの分布を調べ、シーンを分類する。具体的には、まずフレーム中の動きベクトルの分布を計算して、各フレームが図 4 (a) ～ (e) に示す 5 つのタイプのどれに属するかを調べる。

【0033】

(a) フレーム中に動きベクトルがほとんど存在しない（動領域のマクロブロック数が $Mmin$ 以下）。

【 0 0 3 4 】

(b) 同じ向き・大きさの動きベクトルが画面全体に分布している（動領域のマクロブロック数が M_{max} 以上で大きさと方向がある範囲内にある）

(c) フレーム中で特定の部分にだけ動きベクトルが現れている（動領域のマクロブロックの位置が特定の部分に集中している）

(d) フレーム中に放射状に動きベクトルが分布している。

【 0 0 3 5 】

(e) フレーム中の動きベクトルの数が多く、方向も不揃いである。

【 0 0 3 6 】

これら (a) ~ (e) のケースは、いずれも入力画像信号 1 0 0 を得るときに使用したカメラや、撮影された画像内のオブジェクトの動きと密接に関係している。すなわち、(a) ではカメラもオブジェクトもいずれも静止している。(b) はカメラの平行移動。(c) は静止している背景の中でオブジェクトが動いている。(d) ではカメラはズーミングを行っている。(e) ではカメラ、オブジェクトが共に動いている。

【 0 0 3 7 】

以上のように各フレーム毎に分類した結果をシーン毎にまとめ、シーンが図 4 (a) ~ (e) のどのタイプに属するかを判定する。判定されたシーンのタイプと前記算出された特徴量を用いて後述の符号化パラメタ生成部 3 2 で符号化パラメタであるフレームレートとビットレートをシーン毎に決定する。

【 0 0 3 8 】

次に符号化パラメタ生成部 3 2 の個々の処理について詳しく説明する。

【 0 0 3 9 】

＜符号化パラメタ生成部・フレームレート算出＞

符号化パラメタ生成部 3 2 では、まずフレームレートを算出する。

【 0 0 4 0 】

前述の特徴量算出部 3 1 でシーン毎の特徴量の代表値を算出しているものとする。これに対して j 番目のシーンのフレームレート $FR(j)$ を

$$FR(j) = a \times MV_{num_j} + b + W_FR \quad \text{式 (1)}$$

により算出する。 MV_{num_j} は j 番目のシーンの代表値、 a 、 b はユーザが指定したビットレートと画像サイズに関する係数、 W_FR は後述する重みパラメタである。式 (1) は、動きベクトルの代表値 MV_{num_j} が大きくなるほどフレームレート $FR(j)$ が高くなることを意味している。すなわち、動きの大きいシーンほどフレームレートが高くなる。

【0041】

また、動きベクトルの代表値 MV_{num_j} は、前述のフレーム中の動きベクトルの数の他にフレーム中の動きベクトルの大きさの絶対値和、密度なども用いることもある。

【0042】

＜符号化パラメタ生成部・量子化幅の算出＞

各々のシーンに対するフレームレートを算出した後、次に各々のシーンに対する量子化幅を計算する。 j 番目のシーンに対する量子化幅 $QP(j)$ はフレームレート $FR(j)$ と同様、シーンの動きベクトルの代表値 MV_{num_j} を用いて以下の式で算出する。

【0043】

$$QP(j) = c \times MV_{num_j} + b + W_QP \quad \text{式 (2)}$$

ここで、 c 、 d はユーザが指定したビットレートと画像サイズに対する係数、 W_QP は後述する重みパラメタである。式 (2) は動きベクトルの代表値 MV_{num_j} が大きくなるほど量子化幅 $QP(j)$ が大きくなることを意味している。すなわち、動きの大きいシーンほど量子化幅は大きくなり、逆に動きの小さいシーンほど量子化幅は小さくなり画像は鮮明になる。

【0044】

＜符号化パラメタ生成部・フレームレートと量子化幅の修正＞

式 (1)、式 (2) を用いてフレームレートと量子化幅を決める際、＜特徴量算出部・シーンの分類＞の処理で得られたシーンの分類結果（シーンを構成するフレームのタイプ）を用いて式 (1) に重みパラメタ W_FR を、式 (2) に重みパラメタ W_QP を加え、フレームレートと量子化幅の修正を行う。

【0045】

フレーム中に動きベクトルがほとんど存在しない図 4 (a) の場合にはフレームレートを下げて、量子化幅を小さく取る (W_FR 、 W_QP ともに小さくする)。図 4 (b) ではカメラの動きが不自然にならないように、なるべくフレームレートを上げ、量子化幅は大きくする (W_FR 、 W_QP ともに大きくする)。図 4 (c) では動いているオブジェクトの動き、すなわち動きベクトルの大きさが大きい場合にはフレームレートを修正する (W_FR を大きくする)。図 4 (d) ではズームの際にほとんどオブジェクトについては注目されていないと思われることから、量子化幅は大きく取り、フレームレートをできる限り上げる (W_FR を大きくし、 W_QP も大きくする)。図 4 (e) でもフレームレートを上げ量子化幅を大きくとる (W_FR 、 W_QP ともに大きくする)。

【0046】

このようにして設定された重みパラメタ W_FR 、 W_QP をそれぞれ加えることにより、フレームレートと量子化幅の調整を行う。

【0047】

＜符号化パラメタ生成部・マクロブロック毎の量子化幅の設定＞

ユーザからマクロブロック毎に量子化幅を変化するように指定された場合には、フレーム中にモスキートノイズが出やすいと判定されたマクロブロックやテロップ文字のように強いエッジが存在すると判定されたマクロブロックに対して他のマクロブロックよりも量子化幅を小さく設定することで画質改善を図ることもできる。

【0048】

符号化対象フレームに対して、図 5 のようにマクロブロックをさらに 4 つに分けた小ブロック毎に輝度値の分散を計算する。このとき分散が大きい小ブロックと小さい小ブロックが隣り合う場合に量子化幅が大きいとそのマクロブロックではモスキートノイズが発生しやすい。つまり、マクロブロック内でテクスチャが複雑な部分にテクスチャの平坦な部分が隣接する場合にモスキートノイズが出やすくなる。

【0049】

そこで、輝度値の分散が大きい小ブロックに分散が小さい小ブロックが隣接し

ている場合をマクロブロック毎に判定し、モスキートノイズが出やすいと判定されたマクロブロックについては他のマクロブロックよりも相対的に量子化幅を小さくする。逆に、テクスチャが平坦でモスキートノイズが出にくいと判定されたマクロブロックに対しては他のマクロブロックよりも相対的に量子化幅を大きくし、発生符号量の増加を防ぐ。

【 0 0 5 0 】

例えば、j 番目のフレーム内の m 番目のマクロブロックについて、マクロブロック内に小さな 4 つのブロックがあるとき、図 5 に示すように

$$(\text{ブロック } k \text{ の分散}) \geq \text{MBVarThre1}$$

$$\text{かつ } (\text{ブロック } k \text{ に隣接するブロックの分散}) < \text{MBVarThre2} \quad \text{式 (3)}$$

という組み合わせを満たす小さなブロックがあるならば、この m 番目のマクロブロックをモスキートノイズがでやすいマクロブロックであると判定する (MBVarThre1, MBVarThre2 はユーザが定義する閾値)。このような m 番目のマクロブロックに対して

$$QP(j)_{\text{—}m} = QP(j) - q1 \quad \text{式 (4)}$$

のようにマクロブロックの量子化幅 $QP(j)_{\text{—}m}$ を小さくする。これに対してモスキートノイズが出にくいと判定された m' 番目のマクロブロックに対しては

$$QP(j)_{\text{—}m'} = QP(j) - q2 \quad \text{式 (5)}$$

のようにマクロブロックの量子化幅 $QP(j)_{\text{—}m}$ を上げることで、符号量の増加を防ぐ ($q1$, $q2$ は正の数で、 $QP(j) - q1 \geq$ (量子化幅の最小値)、 $QP(j) + q2 \leq$ (量子化幅の最大値) を満たす)。

【 0 0 5 1 】

その際、前述したカメラパラメタの判定で図 4 (b) の平行移動シーン、図 4 (d) のカメラズームのシーンと判定されたシーンについては、カメラの動きに支配されるために画像中のオブジェクトに対する視覚的注目度が低いと思われることから $q1$, $q2$ は小さくとる。逆に、図 4 (a) の静止シーン、図 4 (c) の動いている部分が集中しているシーンでは、画像中のオブジェクトに対する視覚的注目度が高いと思われることから $q1$, $q2$ は大きくとる。

【 0 0 5 2 】

また、文字のようなエッジが存在するマクロブロックについても、量子化幅を小さくすることで文字の部分を明瞭にさせることもできる。フレームの輝度値データに対してエッジ強調フィルタを施し、マクロブロック毎に濃淡値の勾配が強い画素を調べる。画素の位置を集計し、勾配の大きい画素が部分的に集中しているブロックをエッジが存在するマクロブロックであると判断し、式（４）に従いそのブロックについて量子化幅を小さくし、式（５）によりその他のマクロブロックの量子化幅を大きくする。

【 0 0 5 3 】

次にユーザが指定したビットレートを満たすように、上記のように算出された符号化パラメタを修正する符号化パラメタ修正部 3 4 の個々の処理について詳しく説明する。

【 0 0 5 4 】

<符号化パラメタ修正部・発生符号量の予測>

上述のように算出された各シーン毎のフレームレートと量子化幅を用いて符号化すると、シーンのビットレートの割合が許容されるビットレートの上限值あるいは下限値を超える場合がある。そのため、限界値を超えるシーンのパラメタを調整して上限値あるいは下限値内におさめる必要がある。

【 0 0 5 5 】

例えば、前記算出された符号化パラメタのフレームレートと量子化幅で符号化し、ユーザが設定したビットレートに対する各シーンのビットレートの割合を算出したとき、図 6（a）のようにビットレートの上限值あるいは下限値を超えるようなシーンが出てくる場合がある。そこで本発明では、次のようにシーンのビットレートの割合が、許容されるビットレートの上限值あるいは下限値を超えないように修正する処理を施す。

【 0 0 5 6 】

ユーザが設定したビットレートに対する割合を算出したとき、ビットレートの上限值を超えるようなシーンでは図 6（b）のようにビットレートを上限値に設定し直す。同様に下限値を下回るシーンでは図 6（b）のようにビットレートの

割合を下限値に設定し直す。この操作により過剰、あるいは不足となった符号量は、図 6 (c) のように修正しなかった他のシーンに再分配し、全体の符号量を変えないように操作する。

【 0 0 5 7 】

発生符号量は例えば次のように予測する。

【 0 0 5 8 】

各シーンの最初のフレームを I ピクチャ、その他を P ピクチャにすると仮定し、それぞれの符号量を算出する。まず I ピクチャの発生符号量を推定する。I ピクチャの発生符号量については一般的に量子化幅 Q_P と符号量の間に図 7 のような関係があることから、1 フレームあたりの発生符号量 $Code_I$ を例えば次のように算出する。

【 0 0 5 9 】

$$Code_I = I_a \times Q_P^{\wedge} I_b + I_c \quad \text{式 (6)}$$

ここで \wedge はべき乗を示し、 I_a 、 I_b 、 I_c は画像サイズ等により定められる定数とする。

【 0 0 6 0 】

さらに、P ピクチャについては動き補償残差 $MeSad$ と符号量の間にほぼ図 8 のような関係があることから、1 フレームあたりの発生符号量 $Code_P$ を例えば次のように算出する。

【 0 0 6 1 】

$$Code_P = P_a \times MeSad + P_b \quad \text{式 (7)}$$

ここで、 P_a 、 P_b は画像サイズ、量子化幅 Q_P 等により定められる定数とする。画像特徴量算出部において式 (7) に用いる $MeSad$ は計算されているものとし、これらの式から各シーン毎に発生する符号量の割合を算出する。j 番目のシーンの発生符号量は

$$Code(j) = Code_I + (\text{符号化する予定のフレームの } Code_P \text{ の和}) \quad \text{式 (8)}$$

上記の式により算出されたシーン毎の符号量 $Code(j)$ をそのシーンの長さ $T(j)$ で除するとそのシーンの平均ビットレート $BR(j)$ が算出される。

【 0 0 6 2 】

$$B R (j) = C o d e (j) / T (j) \quad \text{式 (9)}$$

このように算出されたビットレートをもとに、符号化パラメタの修正を行う。

【 0 0 6 3 】

また、上記のようなビットレートの修正により予測された符号量を大幅に変更するような場合、各シーンのフレームレートを修正してもよい。すなわち、ビットレートを低くしたシーンではフレームレートも低くし、ビットレートを高めたシーンではフレームレートも高めることにより画質を保つようにする。

【 0 0 6 4 】

以上のように、1パス目で算出されたシーン毎のフレームレートとビットレートを2パス目で符号化部に渡し、動画像信号100を符号化する。符号化部では、1パス目で得られた符号化パラメタ134をもとにシーン毎に目標ビットレートとフレームレートを切り替えながら、従来のレート制御を用いて符号化する。また、1パス目で得られたマクロブロックの情報を用いて、レート制御により算出された量子化幅に対して相対的にマクロブロックの量子化幅を変化させる。これにより、まとまったひとつのシーンの中ではビットレートが保たれるので、符号化されたビットストリームのサイズは目標データサイズを満たすことができる。

【 0 0 6 5 】

図10に本発明の手法と従来法を用いて符号化した際のビットレートとフレームレートの推移の例を示す。従来法（図10（a））では401のように一定の目標ビットレートと403のように一定のフレームレートが設定されている。また、実際のビットレートを402、実際のフレームレートを404で表す。このとき動きの激しいシーンに切り替わると、発生符号量が急増するために前述したように図11のようなフレームスキップがおこり404のようにフレームレートが落ちてしまう。これに対して本発明の手法（図10（b））では、シーンに応じて405のように目標ビットレートを、また、407のように目標フレームレートを定めている。これにより動きの激しいシーンに切り替わったときでもそのシーンに割り当てられたビットレートが多いためにフレームスキップは起こりに

くくなり、フレームレートは目標値を満たすことができる。

【 0 0 6 6 】

なお、本実施形態の符号化装置の機能はすべてコンピュータプログラムによって実現することができ、そのコンピュータプログラムをコンピュータ読み取り可能な記録媒体を通じて通常のコンピュータに導入するだけで本実施形態と同様の効果を得ることが可能となる。この場合のコンピュータの機能構成を図9に示す。

【 0 0 6 7 】

主制御部1はCPUなどから構成されるものであり、各種プログラムを実行する。記憶部2には符号化プログラム6がロードされており、この符号化プログラム6を実行することにより、符号化対象の原画像データファイル7についてのシーン毎の統計的特徴量が得られ、それに基づいてシーン毎に最適化された符号化処理を行うことができる。符号化対象の動画像信号は原画像データファイル7として記憶部2にあらかじめ読み込んでおかななくてもよく、入出力制御部3および入力部4を通じて外部のデジタルVTRやDVDなどから入力される動画像信号を繰り返し入力して2パス符号化を行ってもよい。符号化プログラム6による符号化結果は符号化ビットストリームデータファイル8として出力され、それを出力部5を通じて外部に伝送したり、あるいは復号プログラム9によって復号・再生することができる。

【 0 0 6 8 】

【発明の効果】

以上詳述した如く本発明によれば、動きの大きいシーンでは比較的ビットレートが高く設定されているのでフレームスキップが起こりにくくオブジェクトの動きが滑らかになり、静止したシーンでは画像中の注目されやすい部分であるエッジや文字部分周辺の量子化幅が小さく設定されるので画像が明瞭になる。これにより画像中のオブジェクトの動きを無視した従来の符号化画像に対し、シーン毎にまとまりのある画像が得られ、画質改善効果が得られる。

【 0 0 6 9 】

また、ユーザが設定した上限・下限値内でシーン毎にビットレートを効率よく

配分することで、指定されたデータサイズに合わせたビットストリームを作成できる。さらに、符号化には従来のレート制御を用いることにより、符号化の機構をほとんど変更せずに、目標データサイズを満たすことのできる動画像符号化装置が構成できる。

【図面の簡単な説明】

【図 1】

本発明の一実施形態に係る動画像符号化装置の構成を示すブロック図。

【図 2】

同実施形態における符号化の処理手順を示すフローチャート。

【図 3】

同実施形態におけるシーン分割の処理手順を示す図。

【図 4】

同実施形態における動きベクトルによるフレームのタイプ分けについての説明図。

【図 5】

同実施形態におけるモスキートノイズの発生しやすいマクロブロックの判定について説明する図。

【図 6】

同実施形態における発生符号量調整の処理手順を示す図。

【図 7】

同実施形態における I ピクチャに関する発生符号量を示す図。

【図 8】

同実施形態における P ピクチャに関する発生符号量を示す図。

【図 9】

同実施形態の符号化処理を行うソフトウェアで行う場合におけるコンピュータの機能構成を示す図。

【図 1 0】

同実施形態におけるビットレートとフレームレートの遷移を従来例と対比して示す図。

【図 1 1】

従来法によるバッファとフレームレートとの関係を示す図。

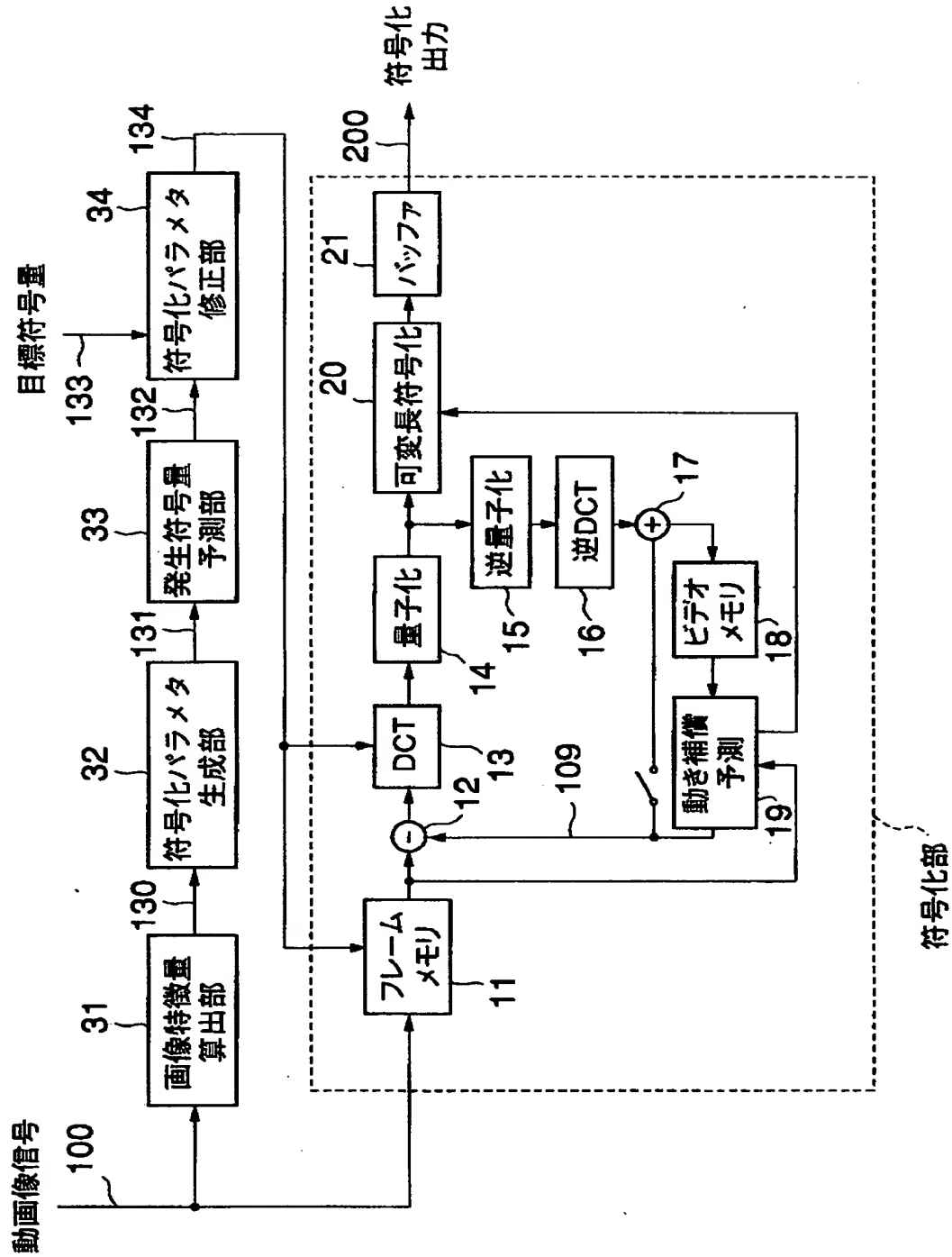
【符号の説明】

- 1 1 … フレームメモリ
- 1 2 … 減算器
- 1 3 … D C T 部
- 1 4 … 量子化部
- 1 5 … 逆量子化部
- 1 6 … 逆 D C T 部
- 1 9 … 動き補償予測部
- 2 0 … 可変長符号化部
- 2 1 … バッファ
- 3 1 … 画像特徴量算出部
- 3 2 … 符号化パラメタ生成部
- 3 3 … 発生符号量予測部
- 3 4 … 符号化パラメタ修正部

【書類名】

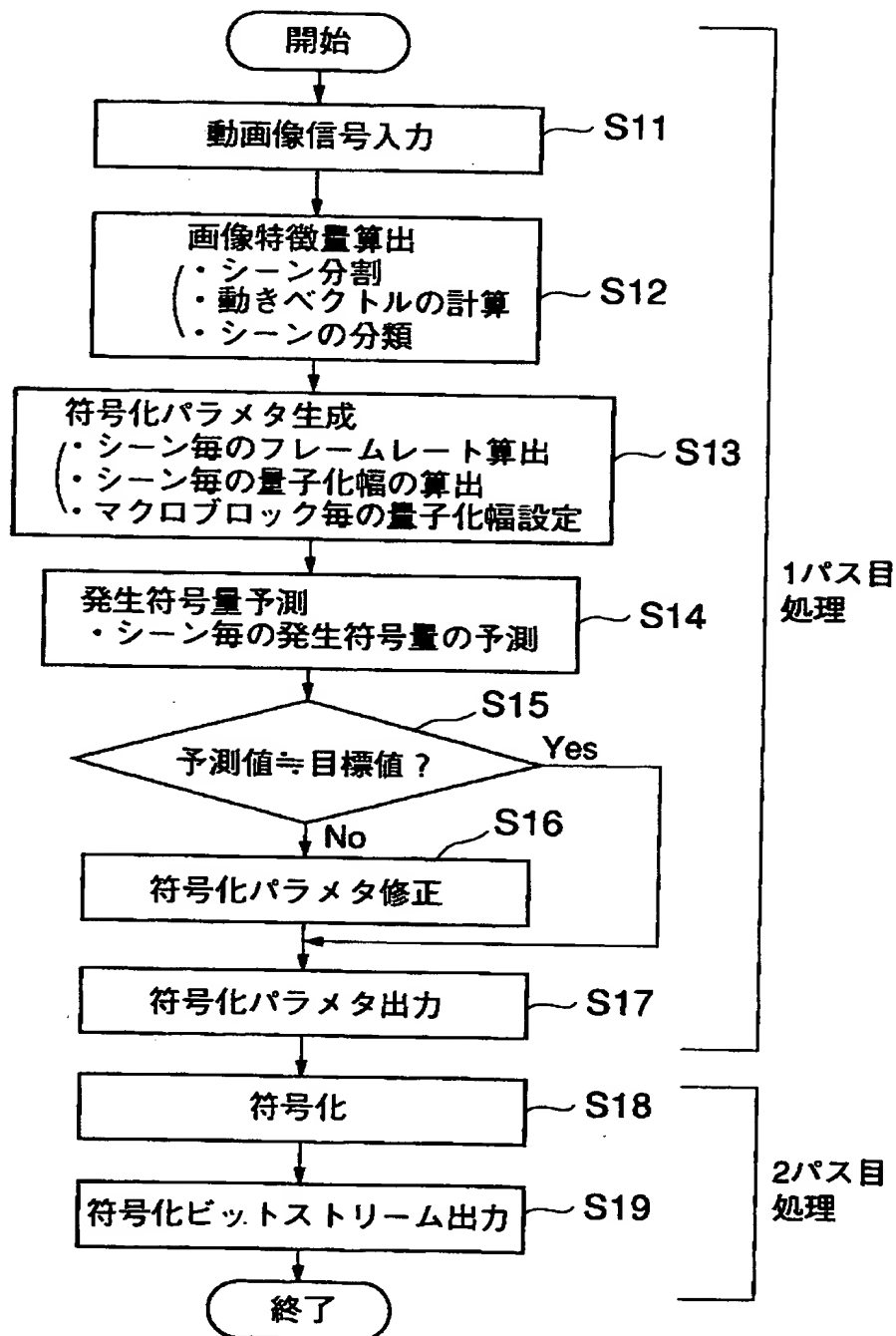
図面

【図 1】



【図 2】

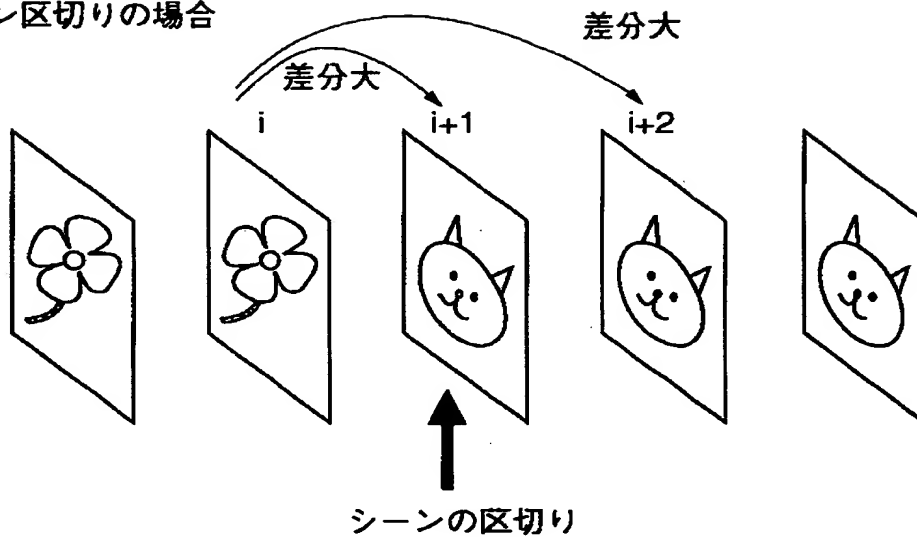
符号化処理の流れ



【図 3】

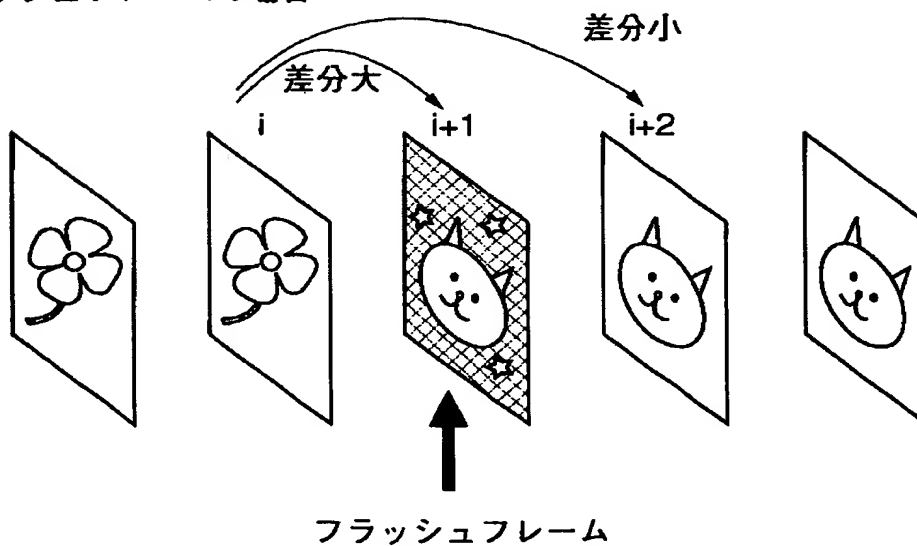
シーン区切りの判定

シーン区切りの場合



(a)

フラッシュフレームの場合

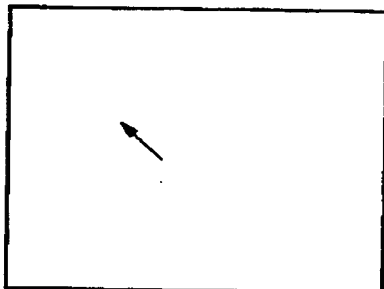


(b)

【図4】

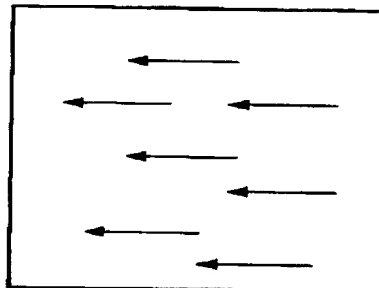
動きベクトルの分類

- ① 動きベクトルがほとんどない場合



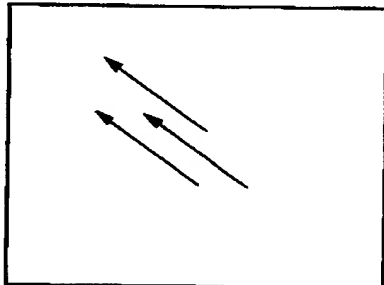
(a)

- ② 同じ向き、大きさの動きベクトルが画面全体に分布している場合



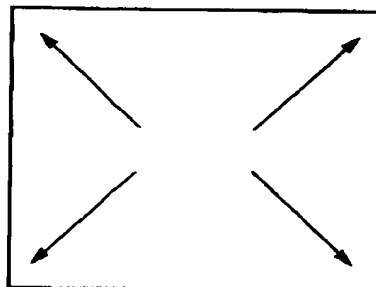
(b)

- ③ フレーム中で特にある部分だけ動きベクトルが現れている場合



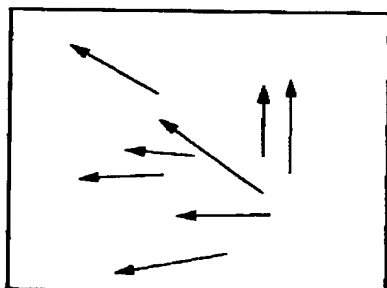
(c)

- ④ フレーム中に放射状に動きベクトルが分布している場合



(d)

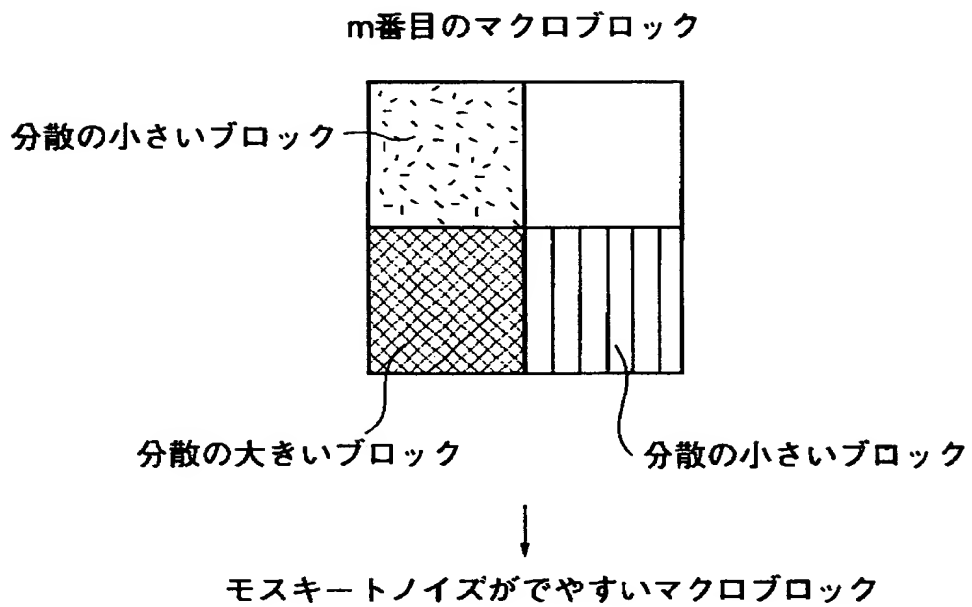
- ⑤ フレーム中の動きベクトルの数が多く、方向もばらばらである場合



(e)

【図 5】

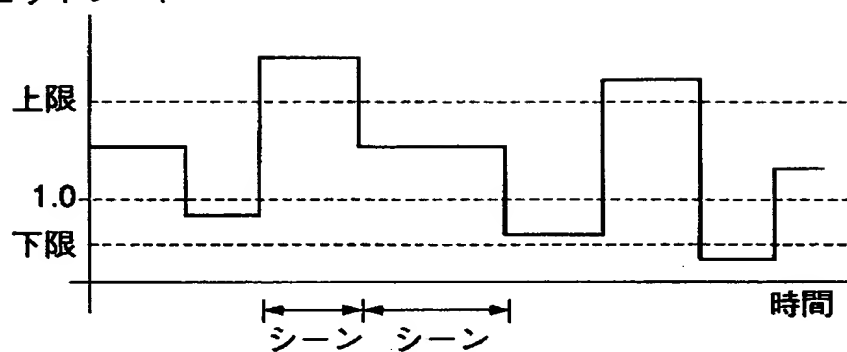
モスキートノイズがしやすいマクロブロックの判定



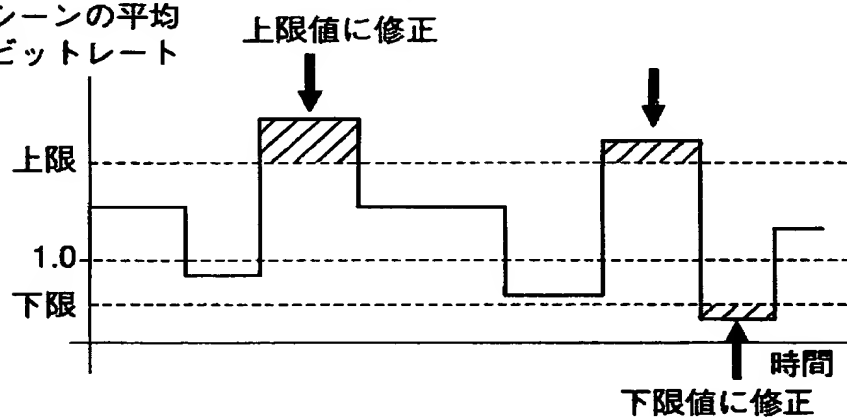
【図 6】

ビットレートの調整

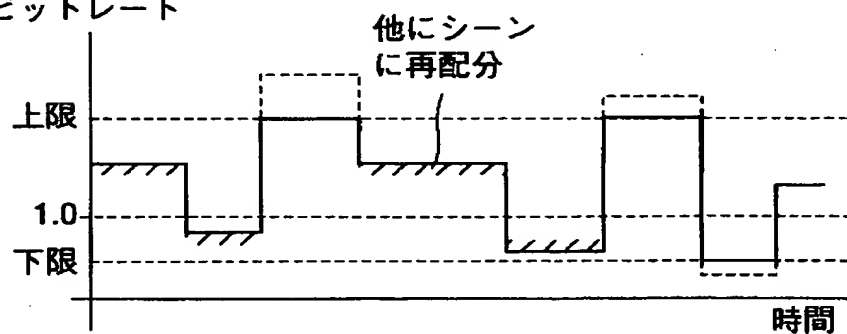
(a) シーンの平均
ビットレート



(b) シーンの平均
ビットレート

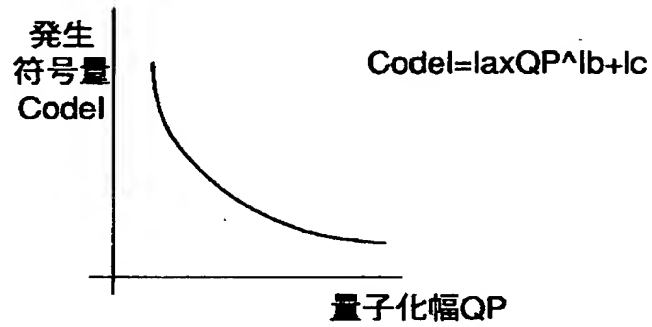


(c) シーンの平均
ビットレート



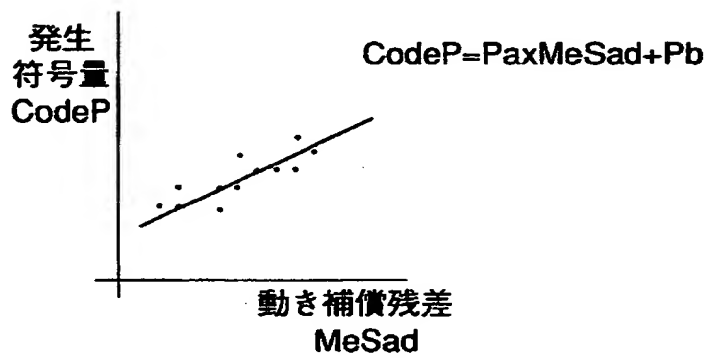
【図 7】

Iピクチャに関する発生符号量

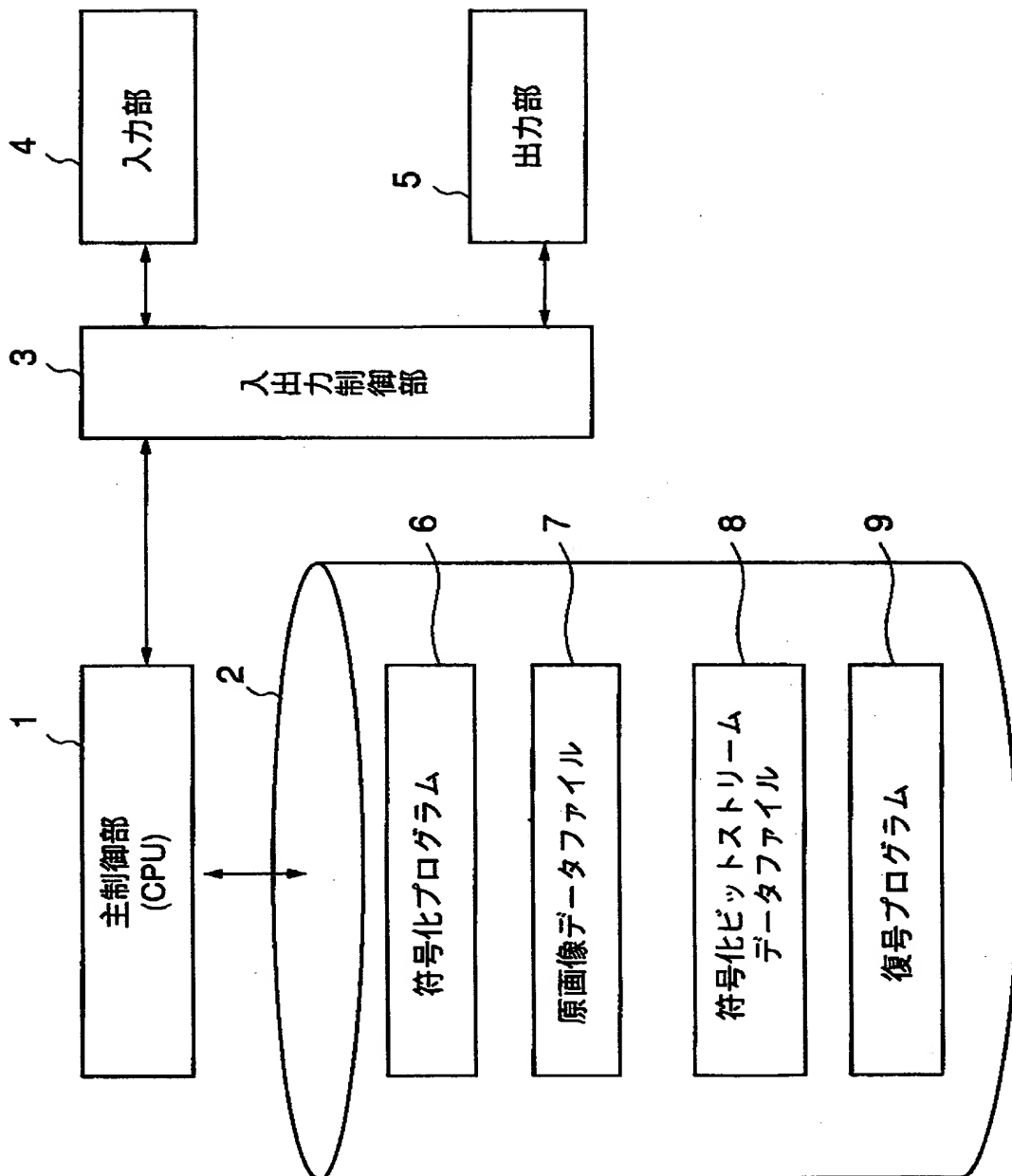


【図 8】

Pピクチャに関する発生符号量



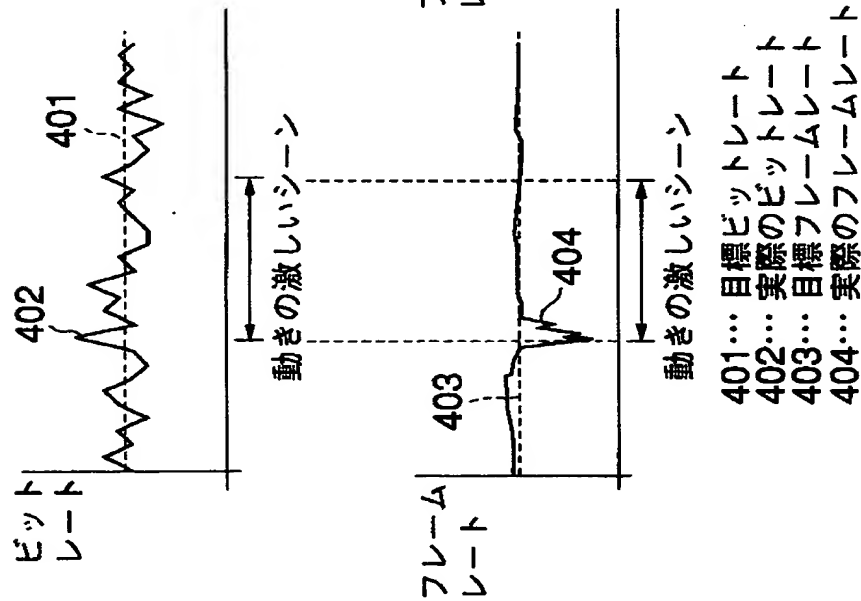
【図 9】



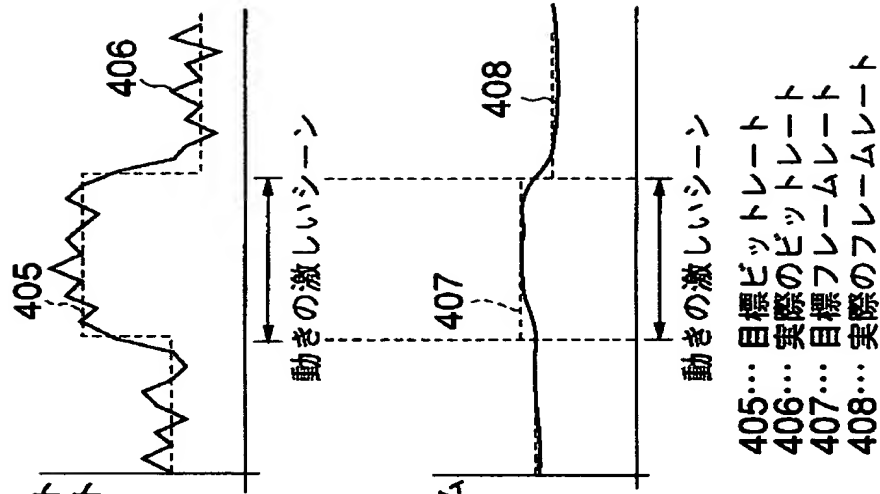
【図 10】

従来法と本発明による符号化画像の
ビットレートとフレームレートの推移

(a) 従来

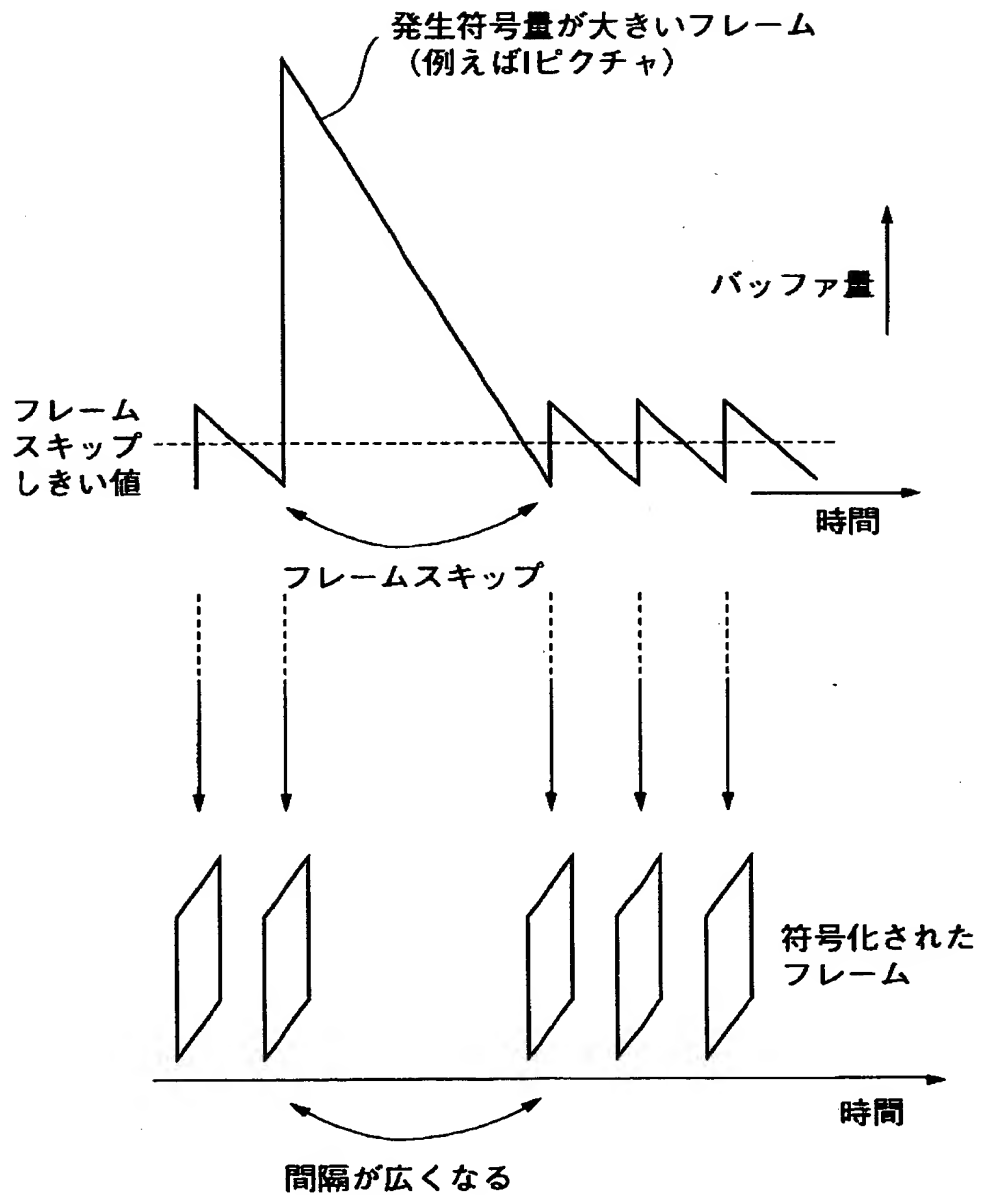


(b) 本発明



【図 11】

従来のバッファとフレームレートの関係



【書類名】 要約書

【要約】

【課題】 符号化する動画像のシーンに応じて符号化パラメタを設定し、シーン内容に適した効率的な符号配分を行い、視覚的にまとまりのある見やすい画像を生成する。

【解決手段】 1 パス目では、画像特徴量算出部 3 1 によって算出された画像の特徴量を基にシーン毎に最適な符号化パラメタを決定し、効率の良い符号量の配分を行う。まず画像特徴量からオブジェクトの動きやカメラパラメタを推定し、その結果を基にシーン毎の符号化パラメタを算出する。オブジェクトの動きの速いシーンではフレームレートを上げて動きが不自然になるのを避け、ノイズがでやすいマクロブロックでは量子化幅を小さく設定することで画質改善を図る。2 パス目では、算出されたシーン毎のフレームレートとビットレート 1 3 4 を用いて、入力動画像信号 1 0 0 に対する符号化が符号化部にて行われる。

【選択図】 図 1

出 願 人 履 歴 情 報

識別番号 [0 0 0 0 0 3 0 7 8]

1. 変更年月日	1 9 9 0 年 8 月 2 2 日
[変更理由]	新規登録
住 所	神奈川県川崎市幸区堀川町 7 2 番地
氏 名	株式会社東芝